DOCUMENT RESUME

ED 257 316

AUTHOP Streiff, Virginia

TITLE Cloze in the Classroom with Notes on Needed Research.

Bilingual Education Paper Series, Vol. 5 No. 3. California State Univ., Los Angeles. Evaluation,

FL 015 068

Dissemination and Assessment Center.

SP., AGENCY Office of Bilingual Education and Minority Languages

Affairs (ED), Washington, DC.

PUB DATE Oct 81

INSTITUTION

NOTE 51p.; Paper presented at the Annual Meeting of

Teachers of English to Speakers of Other Languages

(14th, San Francisco, CA, March 19 J)

PUB TYPE Guides - Classroom Use - Guides (For Teachers) (052)

-- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC03 Plus Postage.

DESCRIPTORS Adult Education; *Cloze Procedure; Elementary

Secondary Education; English; English (Second

Language); Language Teachers; Language Tests; *Native

Language Instruction; Native Speakers; Research

Needs; *Second Language Instruction; Test

Construction; *Testing; *Test Reliability; *Test

Selection

ABSTRACT

The use of the cloze procedure by classroom teachers who are not extensively trained in testing in general or language testing in particular is discussed. Focal points are material selection for cloze; some technical aspects of construction and administration; insights about testing language students naive to cloze; reliability of cloze administered to small groups; and typical native speaker scoring patterns, which may suggest criterion ranges for nonnative speaker cloze score comparison. The information is drawn primarily from the work of graduate students who tested monolingual English speakers, Spanish/English bilinguals, and adult learners of English as a Second Language. General points and things to avoid in material selection for cloze are outlined for grades kindergarten through 6, high school grades, and adult classes. For the analysis of classroom cloze reliabilities, a chart indicates for each grade and section the number of items, the language, the administration mode, the number of subjects, the score type, and the reliability coefficients. Correlations of the cloze procedure with results of other achievement tests are also provided. Information on native speaker cloze scores is also given, including the frequency of word deletion, the language of the test, the test content, level of difficulty, number of items, score type, and mean scores in percentages. (SW)

Reproductions supplied by EDRS are the best that can be made

* from the original document. * * *******************************

October 1981

Bilingual Education PAPER SERIES

Evaluation, Dissemination and Assessment Center California State University, Los Angeles 5151 State University Drive Los Angeles, CA. 90032

CLOZE IN THE CLASSROOM WITH NOTES ON NEEDED RESEARCH

> Virginia Streiff Texas A & M University

> > onginating it.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

EDAC-LA

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION

CENTER (ERIC)

1.: This document has been reproduced as received from the person or organization

ii Minor changes have been made to improve reproduction quality.

 Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

The subject of this publication was supported in whole or in part by the United States Education Department. The opinions expressed herein do not necessarily reflect the position or policy of the United States Education Department; no official endorsement by the United States Education Department should be inferred.

This publication was printed with funds provided by the Bilingual Education Act, Title VII of the Elementary and Secondary Education Act

06 1765, as amended by Public Law 93.380.

CLOZE IN THE CLASSROOM WITH NOTES ON NEEDED RESEARCH*

Virginia Streiff

Interest in cloze as an effective and economical language measure has been increasingly visible in scholarly research.

Studies abound which report the use of cloze as a valid and reliable language proficiency measure, from initial studies (Oller and Conrad, 1971; Oller, 1972; Oller et al., 1973) to more recent ones (Stump, 1978; Streiff, 1978, 1979a, 1979b; Hanzeli, 1977; Hinofotis, 1980). Questions on the effectiveness of certain aspects of cloze have also been raised (Alderson, 1979; Porter, 1978). Oller (1979) recommends cloze as a valuable measure for classroom language testing. However, only sparse information exists about how cloze works for the classroom rather than in research projects. A series of classroom cloze testing projects conducted by ESL and bilingual teachers has yielded valuable practical information for classroom teachers. These projects also suggest additional research topics.

The basic question addressed in this report is: What happens when the cloze procedure is utilized by classroom teachers who are not extensively schooled in testing in general or in language testing in particular? The focal points for discussion are:



^{*}This paper was presented at the 14th Annual TESOL Convention, San Francisco, California, March, 1980.

- material selection for cloze;
- some technical aspects of construction and administration;
- 3. insights about testing language students naive to cloze;
- 4. reliability of cloze administered to small groups; and
- 5. typical native speaker scoring patterns, which may suggest criterion ranges for non-native speaker cloze score comparison.

The findings are drawn primarily from the work of graduate students who participated in a language testing seminar in 1979. The seminar was a cooperative enterprise in which students met for a series of classes, and then pursued projects with their own students. The projects were implemented in classrooms from kindergarten through adult levels among monolingual English speakers, Spanish/English bilinguals, and adult ESL learners. Because we can learn from mistakes, the mistakes as well as the successes will be mentioned. Some insights thus gained obviously lead back to the tenets of standard tests and measurements texts. However, others are specific to use of the cloze procedure. The main findings and recommendations are offered below.

MATERIAL SELECTION

Basically, two types of materials were used. Most were passages drawn from published texts and library books. However, in several instances, teachers followed the suggestion of Streiff (1978) and collected oral child narrative to transform into oral or written cloze passages. Figure 1 presents the main recommendations forthcoming from the teachers' studies. Both the materials



Figure 1

MATERIAL SELECTION FOR CLOZE

Grades K through 6

Avoid: oral interview material that is lengthy and detailed on a single topic for general language proficiency purposes,

published material that is very difficult in readability or very unfamiliar in topic for students, and

stories that are highly idiosyncratic to environments or cultures very different from those of students, especially very unusual names of characters or places.

High School and Adult

Avoid: very juvenile story topics that may be "easy" but an insult to adult learners,

material that is very difficult in readability level, even though students are older.

General Points

- While conducting an oral interview to be used for cloze material, keep topics moving along with questions and statements.
- 2. Use published material with natural breaking points at beginning and end of test material.
- 3. Use material that will be equally familiar or unfamiliar to all students.
- 4. Use instructional material from your own classroom.



that worked well and the problems to avoid are presented below in grade level sequence.

Kindergarten and First Grade

Child narrative refers to child discourse obtained during an oral interview, which is tape recorded, transcribed, then edited lightly by adding a short introductory section and transition sentences between topics. Words are then deleted from the child narrative portions of the final passage. Interview topics that have worked well for cloze material at these grade levels include: family members; the child's favorite activities, foods, and friends; anecdotes about family pets; current and typical school activities; and the child's observations about teachers. Cloze passages constructed from such material have been found to be both comprehensible and interesting to children from Native American, Spanish/English bilingual, and monolingual backgrounds.

Specific themes from published material that have worked well include a reading series story about a child who doesn't smile until "show and tell" time at school because she wants to "show" that she has lost a tooth; a story of a little child who brings each of his favorite objects one at a time to sit with him on his mother's lap; and a fable about an old man, an old woman, and their animal friends who all help pull a giant turnip out of the ground. Highly repetitious tales were particularly effective with kindergarten children.

Two problems were encountered at this level with material selection. One teacher interviewed a kindergarten child for mate-



•

rial and found herself listening to a lengthy and detailed account of a movie the child had seen. She tried her cloze test made of this material with a few other children and found that they quickly grew disinterested in another child's long account of the movie. The teacher then reduced this content and incorporated other topics the boy had mentioned and found much more satisfactory response when she administered the final test. A second problem to avoid is constructing cloze from story material too difficult for the grade level being tested. One teacher constructed two tests, one at first grade readability level and one at the third grade level. S. administered the tests in a first grade other than her own. That classroom's teacher was interested in knowing what the children could do, since she herself was new to this particular classroom. The third grade story was exceedingly difficult for most of the first graders. The seminar student was understandably concerned about the task discouraging the children. However, one good outcome was that two of the children actually did so well on the harder cloze that their own teacher immediately provided them more appropriate material for reading instruction. An alert teacher could administer a series of increasingly more difficult passages to students, allowing them to stop when the work becomes too difficult and determine instructional levels without unduly difficult demands upon the young readers.

Grades Two through Six

Child narrative, as described above, may be used successfully. Additional topics include special skills and interests a



child has developed and special accomplishments such as getting a job or participating in sports events. Especially effective materials were natural science topics and other factual articles from junior level encyclopedias. Topics that posed problems and that should be avoided were stories highly idiosyncratic to environments or cultures very foreign to the students' own and stories containing unusual names. American children, regardless of cultural background, appear to have very similar reading topic preferences, as Streiff (1979b) found in reviewing reading interests of Eskimo and mainstream American children.

High School and Adult Levels

Appealing oral interview material for cloze tests directed at foreign students included the followng: other foreign students' first impressions of this country, problems encountered, and pleasant surprises about the United States; school and leisure activities, and differences between this country and the students' home countries. A pertinent suggestion for teachers who teach content areas other than language to foreign students is to tape record their own teaching presentations and use portions of these recordings as language test material. Such cloze content would reveal quickly how comprehensible the teacher is to the student. Also, portions of television newscasts can be recorded and transcribed for use as cloze tests.

Reader's Digest articles, passages from ESL texts, and foreign language texts have also worked well for older students. Of particular value are articles from elementary and junior high



level encyclopedias. Seminar students found the topics of rivers, electricity, and astronomy from these sources quite effective, especially for older students who were educated in their home languages but who needed easier reading material in English. Neutral topics with about the same degree of familiarity or unfamiliarity for all students were chosen. One teacher, looking for easy material, initially selected a child's book about a froggy woggy and a ducky wucky in a puddle wuddle! He finally selected a more sophisticated topic but at an easy readability level. For these more mature ESL students, the most satisfactory sources of material seemed to be elementary and junior level encyclopedias. The idea of using children's literature should not be discarded, however. For parents who are ESL students, good children's literature could be of great interest and also of practical value for both parent language learners and their children.

General Points about Material Selection

For oral cloze material, the teacher-interviewer should keep the material-gathering interview moving along with questions, encouraging responses, just as a good human interest writer would do for a news report. If one topic becomes too detailed or idiosyncratic, the material can be edited out later as the test is prepared. An alternative to interviews is to tape record conversations among native speakers of the language to be tested. Hughes (1980) reported this technique to yield a successful written cloze test among his subjects. With increasing definition of the importance of speech events and speech acts to language acquisition and



use (Hatch, 1978; Fraser, 1979; Canale and Swain, 1979), additional potential for interview and conversation-based cloze tests is suggested: for example, conversations where apologies are offered, different types of requests are made, and where greetings and leavetakings are occurring in different contexts.

The most effective published material for cloze content had passages with topically natural breaking points at beginning and end. Topics selected should be of equal familiarity, or as close as possible, for the test-takers.

Cloze is adaptable to most kinds of normal classroom instructional materials, from library books to training manuals in technical subjects. A teacher can discover objectively and quickly what materials students can handle and how much they progress. Thus, an obvious source for cloze is on the classroom shelves. However, some teachers may have problems here, as some foreign language teachers discovered when they wanted to develop pragmatic tests for their students: Their texts had no intact passages in them, containing only explanations of grammatical rules, short pattern drills, and vocabulary lists.

Examples of Cloze Passages Used by Seminar Students and Others

Following are some samples taken from cloze tests used at various age/grade levels.

Kindergarten, Oral Cloze (Sabino, 1979)

Once upon a time there lived an old man and an old woman who planted a turnip. The turnip grew and grew. At last it was ready to pull. The old man tugged at the turnip.



pulled and he tugged.// He tugged and// But the
turnip would not come of the ground.// Then the
old man the old woman.// The old man tugged
the turnip.// The old woman tugged the 6
Kindergarten, Oral Cloze (Ruvalcaba, 1979)
Then David talks about his family and what they like. "My
mother's name is Olivia father's name is
Raymond.// The inside of my house white.// My mom
likes owls; boy, she really them.// My mom and dad
and myself go to movies.// Then we all go shopping
in the mall."
Grade 1, Oral Cloze (Correa, 1979)
This is a story about David. He is almost seven. David
likes to talk about animals. In this story, David will tell
us about a classroom project on dinosaurs. "I like dino-
saurs. Today we're going to work on our dinosaur centers
nere at school. There are seven children in each center, and
so we have four dinosaur groups. The teacher will give each
one of us some clay. We will make dinosaurs out of clay.
(In) my group I will work with (my) best friend
ernest. Ernest and I <u>(are)</u> very good friends. He is very
(good) at making dinosaurs out of clay. (I) am very
good at making turtles (out) of clay. Our group plans to
(draw) rivers and volcances on construction caper "



Grade 4, Oral Cloze (Streiff, 1978)

Thomas is just nine years old, but he has a job. In this part of the story, Thomas told about why he got a job, how he got it, and what he does. Thomas decided to look for a job one Saturday when everyone in his family was busy but him. Thomas said, "I just told my mom that I was real ____// and I just told her I want to go _____ a job."// She said, "Go ahead." Thomas said, "She just thought, 'Who's going hire you?""// Thomas rode his bike down to the store. He said, "I asked one of the workers. Then said I could have a job. "// Thomas said, "I came 'What?' And I go, 'I got a job!'"// And _____ goes, "You a job." Grades 3-4, Written Cloze (Ferns and Ball, 1979) Few people ever see a beaver in the wild. And that is what the beaver wants. Millions of beavers have been killed by So the beaver seeks a lonely _____ or pond, away from towns and _____. He generally works and eats at _____ because he does not like to _____ himself in the daylight. Grade 6, Oral Cloze (Streiff, 1978) Then Glenn said, "____ of my friends that are on the track team _____ to my school.// Sometimes I go over to



ir and then we skateboard.// We have a lot of $\frac{36}{\sqrt{}}$,// and sometimes we ride doubles on the skateboard.

We to do tricks."//

Adult, Written Cloze (Autry, 1979)

Electricity is a form of energy. We know that it is by what it does. Electricity (gives) us light, heat, and power.

It brings us radio, television, (motion) pictures, and the telephone.

Electricity comes into our (homes) and schoolrooms on wires. The electric current follows a (wire) circuit, usually made of copper or aluminum wires. Electric (current) is easily turned on or off by switches.

CLOZE CONSTRUCTION AND ADMINISTRATION

Construction

For knowledge of the independent, instructional, and frustrational reading levels of students, the work of Bormuth (1968) and



Peterson et al. (1973) establishes the requirement for deletion of every fifth word for readers from elementary school through college and adult vocational training levels. The latter work re-validated the 57 to 60 percent (exact word score) range as reflecting independent reading capacity.

However, for sampling the general language proficiency (versus reading levels) of learners, the seminar students universally preferred the more lenient schemes of every seventh or eighth word. Haskell's (1973) findings support this preference. The seventh word deletion scheme seems to provide the best balance between test effectiveness on the one hand and minimizing test length and test-taker frustration on the other.

Seminar students recommended using easier rather than more difficult material, particularly for audiences new to cloze testing, for the initial cloze experience. They expressed concern over clarity of test copies; one teacher was dissatisfied with the blurry results produced by the school ditto machine. The result actually amounts to "visual noise" in a cloze test, a topic of interest in language test research but not appreciated by the class-room teacher to whom student reactions are important. Most teachers of young children allowed for their students' large print by allowing extra long lines for blanks and extra space between lines. One who didn't thought it would have been a good idea, since the children's writing overlapped printed words above and to the side of the blanks.

After proofreading their cloze tests, several teachers made these recommendations: (1) leave several sentences intact at the



beginning of the test, and (2) be sure the number of words between blanks is the same throughout the test. The effect of moving a blank to the right or left is negligible (Bormuth, 1964); however, it then becomes tempting to deliberately leave in certain words that the pre-selected deletion scheme would remove. While further cloze studies for researchers may be indicated, it seems wisest for classroom cloze users to follow the established procedures.

Administration

Two interesting ideas that were suggested for kindergarten teachers were use of a flip chart and of a noise signal for blanks during oral cloze testing. The flip chart provided a large visual reproduction of the story with blanks in place for visual reference. Two teachers tried this technique to provide students with added cues during the test. One, working with monolingual English speakers in a private kindergarten, found that the flip chart added nothing to test effectiveness. Another observed that it seemed to help less proficient students but was not needed for her more capable students. This technique appears worthy of further study. The same teacher used a Halloween noisemaker with a clicking sound to signal blanks during the oral cloze, rather than simply saying the word "blank." Again, she noted that it seemed to help the less capable students but not others.

Two of the major points forthcoming from primary and elementary grade projects are interrelated: administering cloze in shorter segments and providing breaks during testing sessions. The suggestion to divide the 50-item cloze tests into two shorter



tests of 25 blanks each was made after several projects. The same teachers concluded that they should have anticipated building in breaks for their students part way through the test. cases, teachers kept the students working for unduly long periods of from 60 to 90 minutes on written cloze tasks. This demand seemed to arise out of teachers' efforts to have children try very hard to supply words even when it was obviously too difficult for them. The teachers were evidently rot convinced in advance that the "skip it" rule built into previous oral cloze testing with young children (Streiff, 1978) had merit. This view may arise from teachers who are accustomed to tests on which the "best" students achieve scores of 100 percent or close to it. This is not typical for cloze performance, particularly among kindergarten children who typically fail to supply function words. For oral cloze, it is strongly recommended that teachers read cloze material alone as a good story, not as a series of isplated test sen-Repeating over and over with a single sentence while children puzzle only makes them lose track of the important overall meaning of the pasage, which is what helps them fill in the words in the first place. Listener motivation declines and the pragmatic value of the test is lost.

Though the points may seem obvious, here are two more recommendations from teachers' work with individual oral cloze: (1) be sure the student has been to the bathroom in advance; and (2) don't administer the test during recess, special events, or children's favorite activities.



Teachers of high school and adult students found that test administration for an initial cloze experience proceeded most smoothly when students were assured they would receive the results of their tests and when students and teacher enjoyed good rapport and mutual respect. Length of test time was not a problem at these levels. In several instances, teachers administered cloze tests more than once. They reported that the second occasion proceeded much more smoothly than the first, after students had received and reviewed initial cloze test results. This finding leads to the next topic, working with students naive to cloze procedure.

TESTING STUDENTS NAIVE TO CLOZE

Several seminar students working with adult ESL students had difficulty eliciting interest in a cloze test. These students seemed more focused on passing one of several well-known multiple-choice ESL tests, which determine whether they can proceed to higher levels of education or training. In these cases, having good rapport with their students helped them get through the first cloze session. The adult ESL students seemed more relaxed and interested in trying the task on the next occasion. In one class, an adult ESL student whose instruction had largely been pattern practice, rule explanation, and vocabulary work commented, "This is really good. We should be doing this at least two or three times a week." Valette (1977) notes that initial negative reaction to cloze does not seem to be universal among older students. Rather, it seems to be present among students whose main task is



to pass multiple-choice ESL examinations used to determine entry and exit status of students. Two examples illustrate this point. Foreign language teachers who initiated cloze in their university classes were surprised at how interested the students were and how hard they worked at it (Lantolf and Streiff, 1980). In one adult ESL situation, 72 students scored significantly better on a multiple-choice form of cloze than on a standard open-ended cloze with passages taken from the same chapter of a non-fiction work (\overline{X} = 34.88 and 11.79, respectively; \underline{t} = 20.97, \underline{p} = .000) (Oller and Streiff, in press). It has been observed that multiple-choice cloze is easier (Hinofotis and Snow, 1980); but in this case, the test-givers observed that the students were also motivated. The students even commented that "circling the letters" was better practice for the tests they would eventually have to take.

Seminar students also found more and less effective ways to handle discussing cloze results with their students. An almost chaotic response occurred when one teacher placed his correctly filled-in cloze model on the overhead projector and returned students' corrected tests to them. Students were understandably curious about their own particular errors, and the teacher concluded that a better approach was needed. The next time, he again used a correct model on the overhead projector but returned a clean copy of the cloze test to each student. This time discussion and questions proceeded in a more systematic manner. Students were then given their corrected tests to compare with the model, and the session was considered very profitable by both teacher and students. Following are suggestions from seminar students who admin-



•

istered cloze to older students who had never before experienced the procedure.

- Use short cloze passages as teaching tools during regular instruction for a period of several weeks or more before administering a cloze test.
- Administer a short preliminary cloze test and then give students a complete copy of the passage to study.
- 3. Correct student work on cloze by using an overhead projector with responses in place, immediately after student independent work.
- 4. Work with cloze regularly as an instructional technique.

Interest in using cloze as a teaching procedure arose from this work with older students as well as from projects involving younger students. Figure 2 summarizes a sequence of steps extrapolated and refined from a cloze teaching procedure reported by Kingston and Weaver (1970). Classroom teachers of young second-language learners have reported that children enjoy the cloze learning tasks. Kingston and Weaver employed similar steps using language experience stories of rural Appalachian primary children and found significant benefits to their reading skills.

RELIABILITIES OF CLOZE AMONG SMALL GROUPS

The classroom groups involved in seminar projects mainly ranged from 5 to 25 students. Several larger group's are also included in this report. Test reliability is a key feature in evaluating language measures. Two mechanical features that contribute to test reliability are larger, as opposed to smaller, numbers of test items and of test takers. While .90 is the reliability criterion generally considered acceptable for standardized tests,



Figure 2

A BASIC CLOZE INSTRUCTIONAL TECHNIQUE

- 1. Write short cloze passages on the board or on an overhead transparency.
- 2. Supply a list of the appropriate words at one side of the board or transparency.
- 3. Let students volunteer to read the passage aloud, selecting the words they think appropriate. Write them in as students give them.
- 4. Provide students a written copy of the same passage and have them fill in their words independently.

Variations

- 1. Don't put any potential word fill-ins on the board or transparency. Rather, allow students to contribute as many words as they can for particular blanks.
- 2. Use students' own discourse as the source of short cloze passages.



Valette (1977) suggests that a good classroom test will have a reliability co-efficient in the range of .60 to .80. This lower range is accounted for partly by the small numbers of test takers usually involved, compared to those upon which standardized tests are normed. Table 1 presents KR-21 reliabilities of the cloze tests employed in the seminar students' projects. These figures are listed along with other pertinent information about the tests: the language of the test, the mode of administration (oral, written), and the scoring procedure (exact, acceptable).

The reliability co-efficients fall within the desirable range for classroom tests and in 64 percent of the cases exceed .80. In the majority of cases, the acceptable word scoring procedure yields the more reliable scores, as Streiff (1979a) and Hinofotis (1980) report. The cloze tests clearly revealed the differences among abilities of students who comprised the test-taking groups at all levels from kindergarten through adult levels. Many of the seminar students observed that the cloze scores accorded well with their students' typical classroom work and grade averages. A summary of the cloze reliabilities by major grade level groups is presented in Table 2.

Some seminar students were able to obtain other kinds of test scores, along with the cloze scores, to analyze for criterion-related validity. Table 3 presents a sample of these correlation findings representing the range of grade levels. In a number of cases, the exact word scores are more substantial correlates of standardized test performance, as Bormuth (1968) found in examining cloze scoring methods and standardized reading tasks. When



Table 1
CLASSROOM CLOZE RELIABILITIES

Teacher/Tester	n of Items	Test Lang.	Admin. Mode	n of Subj.	Score Type	KR-2 Reliabil	
Kindergarten							
Ruvalcaba*	48	Eng. Span.	oral	20	accept.	.71	.82
Sabino	34	Eng.	oral	12	accept. exact	.77	.56
Grades 1-2							
Correa, 1	50	Eng.	oral, writ.	10	exact accept.	.92 .76	
Moreno, 1	50	Eng. Eng.	writ.	8	exact	.62 .90	
Sanchez, 1	50	Eng.	writ.	20	accept. exact	.89	.91
Escalante, 2	50	Eng.	oral,	21	accept.	.97	.99
			writ.		accept.	.99	.95
		Span.	oral		accept. exact	.91	.90
Grades 3-5							
Ferns and Ball	EO	15					
J	50	Eng.	writ.	16	accept. exact	.90	.87
4				23	mult.ch. accept.	.88 .87	.76
					exact mult. ch.	.83	
Taylor, 3-5	50	Eng.	writ.	6	accept.	.86	.73
	4 7	Span.			exact accept. exact	•96	.91

^{*}Random selection of test subjects.



Table 1 (continued)

Teacher/Tester	n of Items	Test Lang.	Admin. Mode	n of Subj.	Score Type	KR-2 Reliabil	
Streiff and Ibarra*							
4	50	Eng.	writ.	23	accept. exact	.91	.82
_		Span.			accept. exact	.90	.77
5		Eng.		26	accept. exact	.92	.87
•		Span.			accept. exact	.90	.85
High School							
Angert, 9-11	50	Eng.	writ.	85	accept.	.70	.65
10-12				45	accept. exact	.74	.63
Brooke and Chernoff	50	Eng.	writ.	49	accept.	. 17	
Daniel	50	Germ.	writ.	16	accept. exact	.80	.62
Adult Classes							
Autry	50	Eng.	writ.	8	accept.	.70	
Powell	49 50	Eng.	writ.	5	accept.	.70 .75	•
Sommer	50	Eng.	writ.	16	accept.	.98	.96
	50				accept. exact	•99	.98
Krueger, Berk, DiQuinzio, Ball, Osborn*	50 50	Eng.	writ.	72	exact mult. ch.	.81 .88	

^{*}Random selection of test subjects.

Table 2

SUMMARY OF CLASSROOM C.JZE RELIABILITIES

Grades K-5	Range of KR-21 Reliabi ities
N of subjects range from 6 to 26	
Acceptable word scoring	.7196
Exact word scoring	.6699
High School	
N of subjects range from 16 to 85	
Acceptable word scoring	.7080
Exact word scoring	.6265
Adult Classes	
N of subjects range from 5 to 72	
Acceptable word scoring	.7099
Exact word scoring	.8198



Table 3

SAMPLE CORRELATIONS OF CLOZE WITH OTHER MEASURES

Kindergarten (Sabino)

	Tests of Basi	Peabody Picture Vocar Mary Test	
Cloze	Language	Concepts	
Exact	. 73 (. 01)	.69 (.01)	.52 (.01)
Accept.	.48 (.05)	.66 (.01)	.48 (.05)
N = 12			

Second Grade (Escalante)

	Cali	est	
Cloze	Total Read.	Total Lang.	Total Batt.
Oral Eng. Exact	.79 (.001)	.71 (.002)	.68 (.001)
Writ. Eng. Accept. $N = 21$.71 (.001)	.72 (.001)	.76 (.001)

Fourth, Fifth Grades (Ibarra)

					Metropolitan Achievement Test				
		Cloze	•		Word	Knowl.	Read. Comp.	Read. Total	
4th N =	23	Writ.	Eng.	Exact	.61	(.01)	.72 (.001)	.70 (.001)	
	2. J	Writ.	Eng.	Accept.	.64	(.001)	.72 (.001)	.71 (.001)	
5th	26	Writ.	Eng.	Exact	.69	(.001)	.58 (.01)	.69 (.001)	
N = 26	Writ.	Eng.	Accept.	.62	(.001)	.58 (.01)	.66 (.001)		

Adult (Krueger et al.)

Correlations of Cloze and English Comprehension Levels with Single Factor

N = 72	A.		Factor	В.		Factor
		Standard Cloze	.85		Mult. Choice Cloze	.82
	,	English Comprehension Levels	.82		English Comprehension Levels	.83



this finding is contrasted with the results of Hinofotis (1980), in which acceptable word scores yielded higher correlations than exact word scores with ESL tests, several points are suggested.

First, there are some obv us differences between the present test situations and those of Hinofotis (1980), i. e., numbers of students in the groups, agas, and educational levels of the students, and particular types of tests (basically, elementary level achievement tests versus adult ESL tests). These different potential sources of variance suggest the next point. Several sources of cloze performance variance have been investigated recently, such as passage-wide constraints versus short-range constraints (Chihara et al., 1977), alternate placement of blanks in equivalent cloze forms (Porter, 1978; Streiff and Angert, in press), and passage readability and distance between blanks (Alderson, However, additional and more comprehensive categories are indicated by the empirically-based model developed by Klare (1976), derived from 36 recent research studies on readability. Klare concentrated on 28 reported variables toward the goal of reassessing the validity of readability formulas. His findings are that reader performance is a function of five major variables interacting with one another: the test situation, reader competence, readability level of the material, content of the material, and reader motivation. Such a model has value for cloze language proficiency studies, indicating predictable sources of variance and additional research topics as well. For example, there are ten variables in this model subsumed under "test situation." These variables comprise the sub-categories of: (1) motivating



factors, (2) test readability in relation to text readability, and (3) test effectiveness factors (e.g., liberal reading time). The point, with respect to exact and acceptable scoring procedures and the varying correlations they produce with criterion measures, is that there are numerous systematic and interacting sources of variance. The remarkable feature of cloze is that under such varying conditions, cloze consistently yields acceptable reliabilities and significant correlations with other measures.

The final point is suggested by considering the higher correlations between cloze and other tests of ESL proficiency when acceptable rather than exact word scoring is used (Oller, 1972; Hinofotis, 1980). Acceptable word cloze appears to yield the more reliable language proficiency measure. However, for reading comprehension correspondences with standardized reading measures, exact word scoring is supported by previous research. These studies have revealed that exact word cloze scores (with an every-fifthword deletion scheme) of 57 to 60 percent and 38 to 44 percent, respectively, correspond to the 90 percent and 75 percent (independent and instructional) reading levels, respectively, on standardized reading measures (Bormuth, 1968; Peterson et al., 1973). For the present, then, a practical recommendation is to obtain reading comprehension information with exact word scores using the fifth-word deletion scheme and to obtain language proficiency information with the acceptable word scoring that seems intuitively appealing to teachers using an every-seventh-word deletion scheme. This suggestion acknowledges that language proficiency and reading ability share a great deal of variance



other research has indicated (Oller and Perkins, 1978; Streiff, 1978, 1979a; Oller and Hinofotis, 1980; Oller and Streiff, in press) but are not identical constructs. One can certainly exhibit language proficiency without having reading ability but not the reverse. Two additional recommendations are pertinent. One is to use oral cloze as a direct measure of oral proficiency and written cloze for reading comprehension. Another suggestion on behalf of non-native speakers, for whom teachers prefer more lenient deletion schemes, is further research into standardized reading test correspondences with more lenient cloze deletion schemes.

A problem remaining for written cloze language proficiency scores, already tentatively solved for reading comprehension scores, is that of criterion levels of acceptable word cloze percentage scores. The following section addresses this topic.

TYPICAL NATIVE SPEAKER CLOZE SCORE LEVELS

A question of interest is: Can one tell by non-native speakers' acceptable word scores how closely they perform to native speaker proficiency levels? Two alternatives have been suggested in previous work with cloze, and a third possibility arises out of the work of the seminar students, along with a review of previous studies.

Approach One

One alternative is proposed by Oller et al. (1973).

Testing Thai, Vietnamese, and English-speaking natives in both



their native languages and the first two groups, then, in their second language, English, they found that careful interpretation of a cloze test from one language to the other yielded two cloze tests of comparable difficulty. By admistering a native language cloze test to students, one could get an idea of how well students could potentially perform in the second language. There are some practical problems with this suggestion. An ESL teacher may not be able to prepare a native language version of the test. Even with native speaker assistance, the problem isn't solved, since native speakers are not automatically equipped simply by virtue of being native speakers to produce careful interpretations so that both tests are of comparable difficulty. There are, of course, other situations in which this procedure may be a ready solution to problems of determining how well learners should ideally be expected to perform in a second language. University foreign language classrooms are one such example.

Approach Two

Another route to having a standard for comparison for nonnative speaker performance and progress toward a reasonable close
test gcal is that employed by Streiff (1978). This approach involves testing a group of native speakers of the same age and/or
grade background as that of a group of non-native speakers with
whom comparison is desirable. In the case mentioned here, school
officials of the non-native-speaking group requested cloze scores
of a "run-of-the-mill middle class Anglo" group as a basis for
comparison with their own students. This approach may have more



problems associated with it than the solution suggested before. Great care is demanded in material selection so that topics will be familiar or unfamiliar on an even basis; locating and obtaining permission to test a criterion group is also riddled with problems. This approach hardly seems practical for classroom teachers, except in instances in which the second- or foreign-language learners attend institutions that also accommodate native speakers of the language.

An Alternative

There may be another route to answer the question of how well one's second- or foreign-language students compare with native speaker performance on cloze measures. A review of average native speaker scores was conducted among the seminar projects and also in several published studies of native speaker cloze testing. average acceptable word scores of native speakers in both groups, seminar projects and published studies, reveal some regularities from the youngest test groups through the oldest. A total of 23 cloze tests, administered from kindergarten through college levels, is represented in the general pattern represented in Figure 3. Figure 4 presents the particular average score ranges found for each major grade level. In several instances, the published studies did not originally present average percentage scores. These percentage figures were derived from the number of cloze items and raw cloze scores reported in the studies (Oller et al., 1973; Oller and Conrad, 1971; Stump, 1978; Streiff, 1978).



Figure 3

GENERAL PATTERN OF NATIVE SPEAKER AVERAGE ACCEPTABLE WORD CLOZE SCORES

Student Levels	Average Cloze Sco	ore Range
	Oral Cloze	Written Cloze
Grades K-2	50%-70%	30%-50%
Grades 3-6	70%-90%	50%-70%
High School, College		70%-90%

Figure 4

AVERAGE SCORE SUMMARIES OF STUDIES REVIEWED

Student Levels	Average Cloze	Score Range
	Oral Cloze	Written Cloze
Grades K-2 7th-10th word deletion	54%-73%	34%-54%
Grades 3-6 7th-10th word deletion	72%-89%	59%-72%
High School, College 6th-10th word deletion		748-928



Acknowledging the limited data available and the varying nature of the tests represented, the patterns appear quite regular enough to merit further attention. On written cloze, average native speaker acceptable word scores increase in approximately 20 percentage point increments from primary grades through middle elementary grades to high school and adult levels. A similar pattern emerges with oral cloze, where even fewer speaker studies are now available. In addition, the oral cloze patterns appear to correspond in a regular way to written cloze patterns, for example, the primary grade average oral score range mirrors the written score range of students of the middle elementary level.

The general patterns on written cloze, 30 to 50 percent for primary grade students, 50 to 70 percent for middle elementary students, and 70 to 90 percent for high school and adult students, reflecting acceptable word scoring, may provide criterion ranges with which non-native cloze performance can be compared. Both patterns, those for written and for oral cloze, are presented as hypotheses, meriting further investigation. Firmer knowledge of this sort could be invaluable to teachers and program administrators in many situations. The present findings emerge from review of a wide variety of studies with many types of content, varying difficulty levels, and several deletion schemes. Considering these facts, the regularities are striking. In general, the above results were achieved under the following conditions:

- acceptable word scoring;
- deletion schemes of every sixth to every tenth word;



- for high school and adult students, written material in the range of approximately fourth to eighth grade difficulty;
- 4. for elementary students, written cloze at the grade level of difficulty corresponding to grades in which students were enrolled; and
- 5. for oral cloze, material constructed from child discourse, with normal features of child language intact (unedited).

Table 4 presents the data particular to each native speaker study or project included in this review. Figure 5 summarizes the average percentage scores of non-native English speakers in the test project groups, for general comparison.

CONCLUSIONS

First, more systematic studies are suggested with cloze; for example, studies that systematically take student motivation into account, reveal correspondences between standardized reading levels and more lenient deletion schemes, student familiarity with topic analyzed in relation to cloze, and reveal oral and written cloze performance among native speakers of different ages. In addition, cloze offers possibilities for imaginative research topics related to different language functions in varying contexts.

Next, several major gains were apparent in the seminar students' testing skills. In brief reports of what they had learned from their own and other students' projects, the students mentioned the following most frequently: recognition of their own contributions to error variance (from the tests themselves, to the test situation, to the test-takers). Specifically, they recognized the value of clear and consistent test instructions in both



Table 4

NATIVE SPEAKER CLOZE SCORES

Teacher/Tester	Word Deletion	Lang. of Subj./Test	Test Content	Test Difficulty	n of Items	Score Type	Mean S	Scores
Kindergarten								
Sabino	every 8th	English	oral, fable: "The Turnip"	low, first, Fry	39	accept. exact	54%	46%
Grade 1								
Streiff (1978)	every 10th	English	oral, child narrative on family, pets, activities		50	accept. exact	61%	31%
	every 7th	English	written, first grade story: "On Mother's Lap"	first grade, Fry	25	accept. exact	34%	24%
Grade 2								
Streiff (1978)	every 10th	English	oral, child narrative on family, pets, school		50	accept. exact	73%	47%
	every 7th	English	written, story: "Andy and the Wild Worm"	second grade, Fry	50	accept. exact	54%	40%

Table 4 (continued)

Teacher/Tester	Word Deletion	Lang. of Subj./Test	Test Content	Test Difficulty	n of Items	Score Type	Mean	Scores
Grade 3								
Ferns and Ball	every 7th	English	written, nat- ural science article on beavers	third grade, Fry	50	accept. exact	64%	36%
Streiff (1978)	every 10th	English	oral, child narrative on family, pets, school	entiment represents	50	accept. exact	72%	44%
	every 5th	English	written, story: "Small Boy Chuku"	third-fourth, Dale-Chall	50	accept. exact	63%	26%
Grade 4								
Ferns and Ball	every 7th	English	written, nat- ural science: beavers	third grade, Fry	50	accept. exact	72%	44%
Streiff (1978)	every 10th	English	oral, child narrative on family, school, child's job		50	accept. exact	82%	45%



Table 4 (continued)

Teacher/Tester	Word Deletion	Lang. of Subj./Test	Test Content	Test Difficulty	n of Items	Score Type	Mean	Scores
	every 5th	English	written, story: "Small Boy Chuku"	third-fourth, Dale-Chall	50	accept. exact	53%	26%
Grade 5								•
Streiff (1978)	every 10th	English	oral, child narrative on family, pets, activities		50	accept. exact	89%	48%
	every 5th	English	written, story: natu- ral science, baby polar bear	fifth-sixth, Dale-Chall	50	accept. exact	66%	44%
Grade 6								
Streiff (1978)	every 10th	English	oral, child narrative on family, school, activities		50	accept. exact	84%	52%
	every 5th	English	written, natu- ral science: baby polar bear	Dale-Chall	50	accept. exact	59%	41%



Table 4 (continued)

Teacher/Tester	Word Deletion	Lang. of Subj./Test	Test Content	Test Difficulty	r. of Items	Score Type	Mean	Scores
High School								
Angert	every 7th	English	written, Reader's Digest: "Joe's Eye Tooth"	sixth grade, Fry	50	accept. exact	74%	63 <i>8</i>
Daniel	every 7th	English	Folk tale: "Till Eulenspiegel"	fourth grade, Dale-Chall	50	accept. exact	93%	72%
Oller et al. (1973)	every 6th	Thai	written, Reader's Digest: type source on Thailand		40	accept.	70%	
	every 6th	Viet- namese	written, Reader's Digest: type source on Vietnam		40	accept.	83%	
	every 6th	English	written, Reader's Digest: "Thomas Edison"		40	accept.	92%	

Table 4 (continued)

()

Teacher/Tester	Word Deletion	Lang. of Subj./Test	Test Content	Test Difficulty	n of Items	Score Type	Mean Scores
College							
Oller and Conrad (1971)	every 7th	English	written, "What is a College?"	∞llege	50	exact	44% (fresh- men) 65% (gradu- ates)
Streiff (1978)	every 10th	English	written, on flying	fourth- fifth	50	accept.	82% 57%



Figure 5

SUMMARY: BILINGUAL AND ESL LEARNERS' ENGLISH CLOZE SCORES

Kindergarten

Exact scoring 28% (oral)

Acceptable scoring ----

Grade 1

Exact scoring 16-52% (written)

Acceptable scoring 24-80% (written)

Grade 2

Exact scoring 44% (written) 49% (oral)

Acceptable scoring 36% (written) 35% (oral)

Grades 3-5

Exact scoring 21-31% (written)

Acceptable scoring 42-58% (written)

High School and Adult

Exact scoring 15-28% (written)

Acceptable scoring 16-36% (written)



spoken and written form, of adequately familiarizing students in advance with the test task, pre-planning reasonable time periods for test administration, and the need for clearly printed test copies and for reasonable test difficulty levels. They also pointed out the merit in using the test results as teaching material. In general, insightful analytical comments too numerous to mention were forthcoming, many reflecting awareness of the interrelationships of test, testing context, and test takers with the reliability and validity of tests.

Finally, some major benefits were apparent with respect to interpreting cloze results. Teachers in bilingual elementary programs observed that cloze was a valuable tool for assessing their students' performance in two languages. Primary teachers as well as teachers of adults became aware of the abilities and weaknesses of their students. Teachers of foreign languages at the high school level said they thought cloze deserved more widespread use in foreign language classrooms. Teachers of adults as well as of younger students saw cloze as a valuable teaching and testing tool. A number of students expressed regret that they didn't have cloze results from earlier in the academic year with which to compare their current results, recognizing the value in pretesting and posttesting, and in periodic progress measurement. One student's statement at the end of a project period represents the typical view. She wrote, after testing first grade children, "I am now in the process of preparing another cloze test, one for a teacher down the hall. I feel I have just encountered something I should have known about long ago."



REFERENCES

- Aldersen, J. Charles. "Scoring Procedures for Use on Cloze Tests," On TESOL '79, eds., Carlos A. Yorio, Kyle Perkins, and Jacquelyn Schachter. Washington, D. C.: Teachers of English to Speakers of Other Languages, 1979, pp. 193-205.
- Bormuth, John R. "The Cloze Readability Procedure," Elementary English, XLV (April, 1968), 429-436.
- . "Mean Word Depth as a Predictor of Comrehension Difficulty," California Journal of Educational Research, XV (1964), 226-231.
- Canale, M., and Merrill Swain. "Communicative Approaches to Second Language Teaching and Testing," Applied Linguistics, I, No. 1 (1979).
- Chihara, Tetsuro, John W. Oller, Jr., Kelley Weaver, and Mary Anne Chavez-Oller. "Are Cloze Items Sensitive to Constraints Across Sentences?" Language Learning, XXVII, No. 1 (June, 1977), 63-69.
- Fraser, B. "On Apologizing." Paper presented at the TESOL Conference, Boston, Massachusetts, 1979.
- Hanzeli, Victor E. "The Effectiveness of Cloze Tests in Measuring the Competence of Students of French in an Academic Setting," French Review, L, No. 6 (May, 1977), 865-874.
- Haskell, J. F. "Refining the Cloze Procedure for English as a Second Language," English Record, XXV (1973), 77-82.
- Hatch, Evelyn. Second Language Acquisition: A Book of Readings. Rowley, Massachusetts: Newbury House Publishers, 1978.
- Hinofotis, Frances Butler. "Cloze as an Alternative Method of ESL Placement and Proficiency Testing," Research in Language Testing, eds., John W. Oller, Jr., and Kyle Perkins. Rowley, Massachusetts: Newbury House, 1980.



- , and Becky Gerlach Snow. "An Alternative Cloze Testing Procedure: Multiple Choice Format," Research in Language Testing, eds., John W. Oller, Jr., and Kyle Perkins. Rowley, Massachusetts: Newbury House, 1980, pp. 129-133.
- Hughes, A. "A Closer Look at Conversational Cloze." Paper presented at the TESOL Colloquium. "The Construct Validation of Oral Proficiency Tests," San Francisco, California, 1980.
- Kingston, Albert J., and Wendell W. Weaver. "Feasibility of Cloze Techniques for Teaching and Evaluating Culturally Disadvantaged Beginning Readers," Journal of Social Psychology, LXXXII (December, 1970), 205-214.
- Klare, George R. "A Second Look at the Validity of Readability Formulas," Journal of Reading Behavior, VIII, No. 2 (Summer, 1976), 129-152.
- Lantolf, J., and Virginia Streiff. "Cloze and the Silent Way."

 Paper presented at the Modern Language Association, Houston,
 Texas, December, 1980.
- Oller, John W., Jr. Language Tests at School: A Pragmatic Approach. London, England: Longman, 1979.
- . "Scoring Methods and Difficulty Levels for Cloze Tests of Proficiency in English as a Second Language," Modern Language Journal, LVI, No. 3 (March, 1972), 151-158.
- proficiency," Language Learning, XXI, No. 2 (December, 1971), 183-195.
- , and Frances Butler Hinofotus. "Two Mutually Exclusive Hypotheses About Second Language Ability: Indivisible or Partially Divisible Competence," Research in Language Testing, eds., John W. Oller, Jr., and Kyle Perkins. Rowley, Massachusetts: Newbury House, 1980, pp. 13-23.
- ______, and K. Perkins. Language in Education: Testing the Tests. Rowley, Massachusetts: Newbury House, 1978.
- , and Virginia Streiff. The Language Factor: More Tests of Tests. Rowley, Massachusetts: Newbury House, in press.



- "J. Donald Bowen, Ton That Dien, and Victor W. Mason. "Cloze Tests in English, Thai, and Vietnamese: Native and Non-Native Performance," Language Learning, XXII, No. 1 (June, 1972), 1-13.
- Peterson, J., E. Paradis, and N. Peters. "Revalidation of the Cloze Procedure as a Measure of the Instructional Level for High School Students," Diversity in Mature Reading. 22nd Yearbook of the National Reading Conference. Boone, North Carolina: National Reading Conference, 1973, pp. 144-149.
- Porter, Don: "Cloze Procedure and Equivalence," Language Learning, XXVIII, No. 2 (December, 1978), 333-341.
- Streiff, Virginia. "Language and Achievement Test Scores Among Bilingual Children." Paper presented at the TESOL Colloquium, "Bias in Standardized Testing: The Language Factor," Boston, Massachusetts, 1979a.
- Reading, Comprehension and Language Proficiency Among Eskimo Children: Psychological, Linguistic, and Educational Considerations. New York: Arno Press, 1979b.
- Achievement Test Scores in a Bilingual Setting, "Language in Education: Testing the Tests, eds., John W. Oller, Jr., and Kyle Perkins. Rowley, Massachusetts: Newbury House, 1978.
- _____, and D. Angert. Are Alternate Forms of Cloze Equivalent? In press.
- Stump, T. A. "Cloze and Dictation Tasks as Predictors of Intelligence and Achievement Scores," Language in Education:
 Testing the Tests, eds., John W. Oller, Jr., and Kyle Perkins. Rowley, Massachusetts: Newbury House, 1978.
- Valette, Rebecca M. Modern Language Testing. New York: Harcourt Brace Jovanovich, 1977.



Appendix

STUDENT PROJECTS

The seminar students whose interest, insights, and dedication to their profession contributed greatly to this report are listed below. They continue to contribute in their chosen professional areas: public schools, federal language training institutions, and adult ESL programs in colleges and universities.

- Angert, D. "Cloze Testing in the High School Classroom."
- Autry, M. "A Comparison of Written Cloze with a Multiple Choice Language Comprehension Test."
- Brooke, V., and E. Chernoff. "Cloze as a Measure among Special Education Students."
- Correa, M. "Oral and Written Cloze among First Graders in a Bilingual Program."
- Daniel, A. "Cloze Testing in the Foreign Language Classroom.
- Escalante, T. "Cloze Testing among Spanish-English Bilinguals."
- Ferns, S., and J. Ball "Standard and Multiple Choice Cloze Among English-speaking Third and Fourth Grade Students."
- Ibarra, A. "Bilingual Cloze Performance in Elementary School." (with Streiff)
 - Krueger, A., S. DiQuinzio, V. Berk, J. Osborn, and J. Ball. "Standard and Multiple Choice Cloze among Adult Foreign Trainees." (with Streiff)
 - Moreno, H. "Cloze Reading Scores of First Graders."
 - Powell, L. "Cloze Testing the Language Proficiency of Adult ESL Students."
 - Ruvalcaba, G. "Oral Cloze Tests of English and Spanish in Kinder-garten."
 - Sabino, M. "Oral Cloze at the Kindergarten Level."
 - Sanchez, S. "Cloze Performance of First Grade Spanish-English Bilinguals."



Appendix (continued)

Sommer, F. "Cloze Testing: A Project with Adult ESL Students."

Taylor, P. "Spanish and English Cloze among Bilingual Elementary Students."



Virginia Streiff has a PhD from Ohio University in psychology and linguistics. She taught graduate-level courses in crosscultural pedagogy and testing, was a guest professor at Cornell University in the Summer Linguistics Institute (1981), and is now at Texas A & M University. Her research and publications span age groups from early childhood to adult bilinguals and second language learners of English as well as other languages. She has also presented papers at numerous TESOL international conventions on language testing.

Among Dr. Streiff's recent publications are: The Language Factor: More Tests of Tests, co-authored with John W. Oller, Newbury House, in press; "The Roles of Language in Educational Testing," Issues in Language Testing Research, ed., John W. Oller, Newbury House, in press; Cloze in the Class-room, Newbury House, in preparation; Reading Comprehension and Language Proficiency Among Eskimo Children, Arno Press Series on Bilingualism, 1978; and "Question-Generation by First-Graders," Bilingualism in Early Childhood, eds., W. Mackey and Theodore Andersson, Newbury House, 1978.

